

# COUNTY-SPECIFIC, REAL-TIME PROJECTION OF THE EFFECT OF BUSINESS CLOSURES ON THE COVID-19 PANDEMIC

A PREPRINT

**Dominic Yurk**

Department of Electrical Engineering  
California Institute of Technology  
Pasadena, CA 91125  
dyurk@caltech.edu

**Yaser Abu-Mostafa**

Department of Electrical Engineering  
California Institute of Technology  
Pasadena, CA 91125

February 11, 2021

## ABSTRACT

Public health policies such as business closures have been one of our most effective tools in slowing the spread of COVID-19, but they also impose costs. This has created demand from policy makers for models which can predict when and where such policies will be most effective to head off a surge and where they could safely be loosened. No current model combines data-driven, real-time policy effect predictions with county-level granularity. We present a neural net-based model for predicting the effect of business closures or re-openings on the COVID-19 time-varying reproduction number  $R_t$  in real time for every county in California. When trained on data from May through September the model accurately captured relative county dynamics during the October/November California COVID-19 surge ( $r^2 = 0.76$ ), indicating robust out-of-sample performance. To showcase the model's potential utility we present a case study of various counties in mid-October. Even when counties imposed similar restrictions at the time, our model successfully distinguished counties in need of drastic and immediate action to head off a surge from counties in less dire need of intervention. While this study focuses on business closures in California, the presented model architecture could be applied to other policies around world.

**Keywords** data-driven · model · COVID-19 · policy · neural-net

## 1 Introduction

From the beginning of the COVID-19 pandemic it was recognized that public health policy would be our most effective tool in limiting the spread of the disease. However, while some measures such as social distancing were quickly and widely agreed upon, other policies such as business closures have proven far more contentious. While the economic impact of business closures is immediately evident, the corresponding amount of public health benefit cannot be directly calculated. In order to make informed decisions policy makers have turned to model-based analyses which attempt to estimate how effective various policies are at slowing the spread of COVID-19. Early in the pandemic, the lack of information on the spread of this new disease made truly data-driven analyses of policy effectiveness impossible. Attempts to model policy impacts were forced to rely on assumptions to make up for the scarcity of available data, using techniques such as curve fitting and SEIR models that impose constraints on disease evolution [1–6]. Unfortunately the pandemic has continued to spread widely over the past year, leading to multiple rounds of tightening, loosening, and re-tightening of public health policies around the world. The variation in timing and austerity of these policies between and within countries has created a variety of natural experiments, allowing researchers to model policy effectiveness in a more data-driven way.

To date, most data-driven policy analyses have been retrospective and focused on overall effectiveness across a number of counties, states, or countries [7–10]. These analyses can be broadly useful for policy makers deciding which controls to apply when attempting to contain an outbreak already in progress. However, such models are unable to identify

how policy effects differ in different areas, overlooking the fact that optimal policy may vary greatly between urban centers and rural communities or between an area with one daily case per 10,000 residents and an area with 100. Furthermore, due to long delays between initial infection and subsequent transmission, testing, and reporting, by the time a COVID-19 outbreak becomes readily apparent even an immediate and strict lockdown will take substantial time to contain it. For example, in the United States the strictest lockdown imposed to date was in New York City on March 22nd, but average daily cases did not peak until almost 3 weeks later on April 11th [11]. Thus, there is a need for a real-time, locality-specific model of policy effectiveness capable of alerting officials to the need for action before an outbreak is well underway.

An attempt to fill this gap has recently been presented which projects future mortality rates in various countries under different policy scenarios using Gaussian processes built on top of an SEIR model [12]. However, the only policy variable used is a "stringency index" which clusters many policies from business and school closures to mask wearing and international travel restrictions across an entire country [13]. This not only obscures the effect of individual policies, but would also be of limited use in a country like the United States where policy varies widely between localities. More detailed analysis has largely been frustrated by the lack of datasets capturing county-level public health policies. This problem has recently been addressed by the Stanford COVID-19 simulation group (SC-COSMO), which released a detailed dataset covering county-level public health policies across California [14]. We present a new model based on this data which uses monotonic neural networks to perform real-time county-level policy effect forecasting in California. In this study we focus on the effect of business closures and present both a test data set supporting model accuracy and a case study demonstrating potential utility to policy makers. The structure of our model is simple and relies on only a few input parameters, making it potentially applicable to other policies and other localities across the US and the world.

## 2 Methods

### 2.1 Input and Output Parameters

#### 2.1.1 Measuring $R_t$

The model presented here focuses on the time-varying reproduction number  $R_t$  as the metric for assessing the state of the COVID-19 pandemic.  $R_t$  is a measure of how many new people an infected individual will transmit the virus to given that the individual was infected on day  $t$ . The primary advantage of this metric is that it is directly impacted by policy changes in real time, compared to other metrics such as COVID-19 cases and deaths which can take weeks to respond to policy changes. Another advantage of  $R_t$  is that it is easily interpretable at a glance; any value above 1 indicates accelerating spread, while any value below 1 indicates decelerating spread. This makes it an ideal actionable parameter for policy makers. Thus, we chose to make the output target of our model on any given day the change in  $R_t$  from 7 days before that date to 7 days after, referred to hereafter as  $\Delta R_t$  or  $\text{del\_}R_t$ .

Because  $R_t$  cannot be measured directly, we used the Rt.live methodology to extract it from testing data [15]. While Rt.live only publishes  $R_t$  values at the state and national levels, its computations rely only on daily positive and negative testing volumes in each region, allowing for universal applicability provided appropriate data. We adapted their code to compute  $R_t$  values for each county in California based on county-level testing data gathered from COVID ActNow [16]. For the 15 out of 58 counties in which negative testing volumes were not reported on all relevant dates, the counties' test positivity rates were assumed to match that of the state overall.

#### 2.1.2 Policy Data

The SC-COSMO public health dataset for California tracks seven different policy categories by county with an austerity score of 0 to 10 on each day, as listed in Table 1 (an eighth policy category, `order_gath`, was excluded because it is not tracked for many counties in the dataset). Each of these order types has a potential impact on the pandemic, and ideally our analysis would have incorporated many or all of them. However, to analyze the effect of a given policy in a data-driven way we needed many instances of the policy loosening and tightening across different counties and times. Only `order_closure`, which on average changed over 7 times per county, definitively met this criterion. Other orders with potentially enough data for analysis were `order_shome` and `order_bubble`. However, these policies rely on individual cooperation and are very difficult to enforce, resulting in inconsistent effectiveness. For example, in the two weeks surrounding the March stay-at-home order in Los Angeles median traffic to recreation and retail businesses fell from 92% of normal to 57% of normal, while the November stay-at-home order corresponded to a drop from 73% or normal to 67% of normal [17]. For these reasons we chose to focus exclusively on `order_closure` for this analysis.

Order	Description	Total Changes	Tighter Changes	Looser Changes
closure	Overall closure status	408	169	239
bubble	Social bubble restrictions	134	106	28
masks	Mask restrictions	112	109	3
shome	Stay at home restrictions	84	52	32
quarres	Resident quarantine restrictions	32	21	11
lab	Lab & tracing requirements	30	23	7
quarvis	Visitor quarantine restrictions	2	1	1

Table 1: Number of County-Level Policy Changes in the SC-COSMO Dataset (May 1st - Dec 1st)

### 2.1.3 Model Inputs

Our model was constructed with 6 input variables: 1) population density, 2) number of new cases per 10,000 residents in the last 2 weeks, 3)  $R_t$  value 21 days ago, 4) difference between  $R_t$  values 21 and 28 days ago, 5) current value of order\_closure, and 6) difference in order\_closure between today and 7 days ago. Parameter 6 is intended to be the operative variable for policy makers, as it allows them to project the effects of potential policy decisions if they were made today. In parameters 3 and 4, a 3 week delay is inserted in order to simulate the "fog of war" present in real policy-making situations; on any given day the real-time  $R_t$  live estimates have large error bars due to delays in transmission, testing, and reporting. As we look backwards these error bars become narrower due to increased data availability up to roughly 3 weeks in the past, at which point the uncertainty settles to a steady-state minimum. Thus, all parameters were chosen such that they would be available to users in real time with a high degree of confidence.

When constructing our model we considered data from May 1st, 2020 to December 1st, 2020. The May 1st cutoff was applied to exclude the early phases of the pandemic when testing was scarce, much of the general public was panicked, and many restrictive policies were enacted at once, all of which obscured individual policy effects. The Dec 1st cutoff was applied to provide a 4 week buffer prior to model development at the end of December, equivalent to the 3 week  $R_t$  input delay plus one week for determining the subsequent change in  $R_t$ . Data points consisted of one snapshot per week of every county in California during this period (excluding those with fewer than 10 total cases in the previous week), for a total of 1387 data points. These data were divided into points before and after October 1st, and the latter points were held in reserve as an uncontaminated test set which was not used until after model optimization was performed (see results and discussion). This was done to ensure that we could obtain as clear of a picture as possible of model performance now and in the future when we are only able to train on data from the past.

## 2.2 Monotonic Neural Network

Over the past decade neural network models have achieved world-leading accuracy in tasks ranging from text, speech, and image recognition to predicting protein structure. This is due to the models' ability to identify hidden patterns in data sets without researcher-imposed assumptions or constraints. In data-poor scenarios this lack of constraints can lead to poor extrapolation and occasionally nonsensical predictions when compared to SEIR models traditionally favored by epidemiologists. However, the COVID-19 pandemic has produced a body of data larger than any previously seen in epidemiology, allowing neural networks to attain high levels of performance in COVID-19 prediction first as supplements to SEIR models [18–20] and more recently as standalone models [21,22]. These results encouraged us to pursue neural networks for the application presented here.

By machine learning standards, our data set of 1387 points is still relatively small, leading to a risk of overfitting. We mitigated this risk in two ways; first by limiting our inputs to the six parameters described above, and second by applying a monotonicity constraint to the output relative to change in order\_closure. Monotonicity constraints have previously been shown to improve accuracy and robustness across a variety of machine learning models, particularly when trained on small data sets [23–26]. In our case the monotonicity assumption seems very applicable; it merely specifies that implementing a tighter policy today will never make  $\Delta R_t$  worse than it would have been under a looser policy, but does not apply any constraints on the magnitude of the difference in  $\Delta R_t$  between policies. We utilized an existing framework which allowed us to construct a standard dense neural network model with a monotonicity constraint applied to only one input [27]. We applied rectified linear unit activation between hidden layers and no activation on the output layer, removing any potential bias towards a sigmoidal output shape.

For model hyperparameter optimization we utilized data from May 1st through October 1st (943 total data points) and performed 5-fold cross-validation to estimate performance. We did not follow the standard practice of selecting a random subset of points for validation, as this would lead to very high correlations between some training and validation points; for example, a model trained on Orange County data from June 1st and 15th should do a very good job predicting

Orange County's behavior on June 8th. Instead, at each step we selected a random subset of 20% of all counties and used all data points from those counties for validation. The error function being minimized was RMS error between predicted and actual  $\Delta R_t$ , training was done with the Adam optimizer [28], and model parameters being optimized were number of hidden layers, hidden layer size, optimizer learning rate, optimizer weight decay factor, and number of training epochs.

### 3 Results

The parameters minimizing validation error were 3 hidden layers of 6 nodes each, a learning rate of  $10^{-3}$ , a weight decay factor of  $10^{-5}$ , and 20 training epochs. This yielded training and validation RMS errors of 0.0403 and 0.0445 respectively. Figure 1a shows an example result from one subset of validation counties, demonstrating that model predictions on training and validation data share similar distributions and similar correlations with true values.

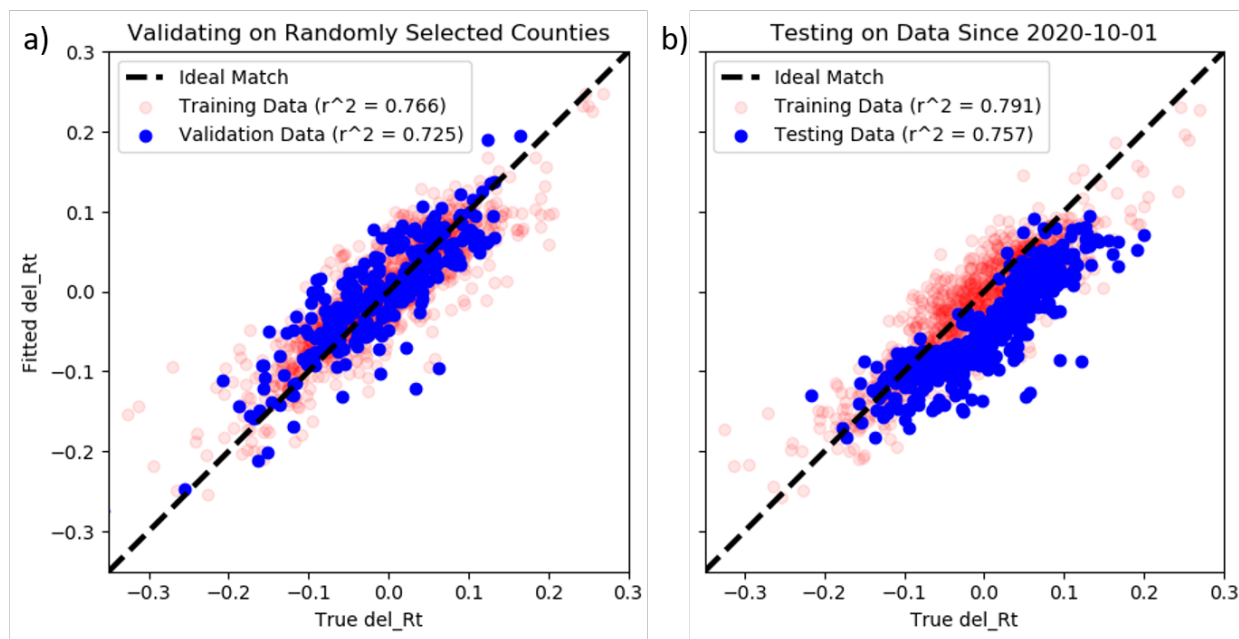


Figure 1: **Model Validation and Testing.** a) To optimize our model we validated on randomly withheld county data from May 1st through Oct 1st and chose model parameters to minimize validation error. b) We tested our model on data from Oct 2nd through Dec 1st. The resulting model outputs have a slight underprediction bias but still capture the overall trend of which counties have relatively good or bad  $R_t$  trends.

After completing hyperparameter optimization we trained our model on the full data set from May 1st to October 1st and tested its performance on our reserved data from October 2nd through December 1st (444 data points). This resulted in training and testing errors of 0.0411 and 0.0625 respectively. At first glance, these error values suggest that our model failed to generalize well to newer data. However, a closer look at all training and testing points in Figure 1b shows that correlation between predicted and true values is almost as good for the test set ( $r^2 = 0.757$ ) as it is for the training set ( $r^2 = 0.791$ ). The problem is that our model is biased towards underpredicting  $\Delta R_t$  in the test set, particularly when the true  $\Delta R_t$  is positive. This mirrors the fact that the ensemble of all major models tracked by the Reich Lab COVID-19 Forecasting Hub [29] consistently underpredicted future incident cases in California for epidemic weeks 41-50 (October 4th through December 6th), coinciding with by far the largest COVID-19 surge in California to date. The reasons for this surge are not yet well understood, but the consistent underprediction shows that it was driven by an unexpected factor that was not captured by almost any of the sophisticated models on the forecasting hub. Our model is relatively simple by design to avoid overfitting, so it should not be surprising that it is also unable to predict the full extent of this surge. However, the strong correlation between our model predictions and true values indicates that it could still provide valuable insights into the relative behavior of different regions under different policy options during such a period by performing county-specific policy impact predictions.

## 4 Discussion

The true utility of our model is its power to project the effects of a range of different policy options in real time for any county. Some randomly selected examples of these projections are shown in Figure 2. It is important to note that our model architecture requires these curves to be monotonically non-increasing, but it does not place any constraints on smoothness or shape. Thus, the observed smooth sigmoidal curves are an emergent property of the trained system. This matches our intuition for how the system should work; incremental policy changes produce incrementally smooth effects, and massive policy changes exhibit a law of diminishing returns. The diversity of centers and ranges in the sampled curves also demonstrates that the same policy change can have starkly different effects depending on when and where it is implemented.

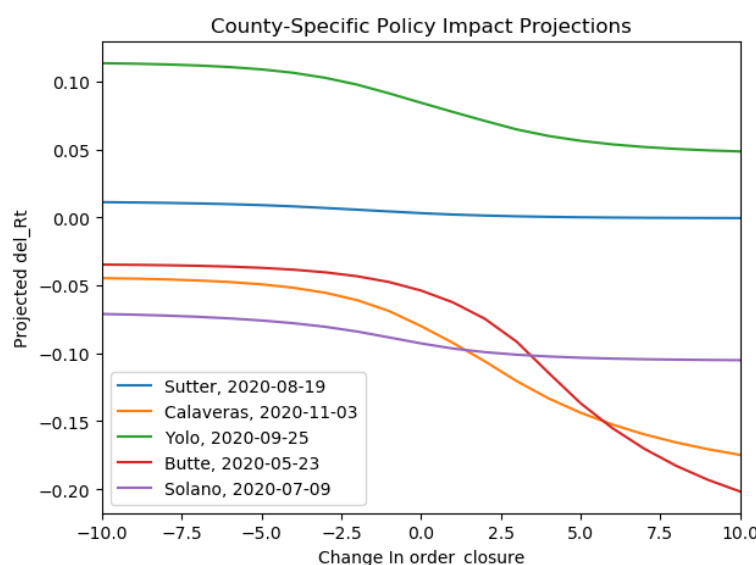


Figure 2: **A Sample of County-Specific Policy Impact Predictions.** We randomly selected five counties and times and show model projections of  $\Delta R_t$  in each case under a range of possible policy changes. This shows that our model produces smooth sigmoid-shaped curves, in line with our intuition.

To demonstrate the potential power of this model we focus on the week of October 15th, which we now know was in the middle of a crucial period of rising  $R_t$  that later produced California’s devastating winter surge. At the time there was relatively little concern over COVID-19 trends; during this week eight counties loosened business closure restrictions and none tightened them. Using only data that were available at the time, our model would have strongly urged some of these counties to change course. We can broadly group the counties into three categories based on model projections; 1) those with a rapidly rising  $R_t$  trend which could be slowed but not reversed by business closures, 2) those with a rapidly declining  $R_t$  trend which would not reverse even with business reopenings, and 3) those at an inflection point where  $R_t$  could rise or fall depending on policy decisions. Figure 3 highlights one representative of each category in the form of Fresno, Mono, and Placer counties respectively. All three counties had relatively loose business closure policies for all of October; Fresno was at 3, Mono was at 2, and Placer dropped from 3 to 2 early in the month. Our model would have suggested that such policy was appropriate in Mono, which showed a substantial decrease in  $R_t$  after October 15th. In contrast, Fresno was at the opposite end of the spectrum during that same week with a rapidly rising  $R_t$ . Our model would have projected that business closures would help slow this trend, but more restrictive policies on top of that would have been necessary to reverse it. Placer was in between these extremes, with our model suggesting a plateau around October 15th under status quo or the potential for substantial reduction in  $R_t$  under business closures. The underprediction bias mentioned earlier is evident here, as Placer did maintain status quo and  $R_t$  continued to rise slowly after October 15th. However, the overall trend was still captured; our model forecast was most dire in the areas which turned out to have the worst  $R_t$  trend. Policy makers at the time could have used these projections to focus business closures on areas like Fresno, thereby heading off surges in incipient hotspots while reducing economic impact in areas with less worrying trends.



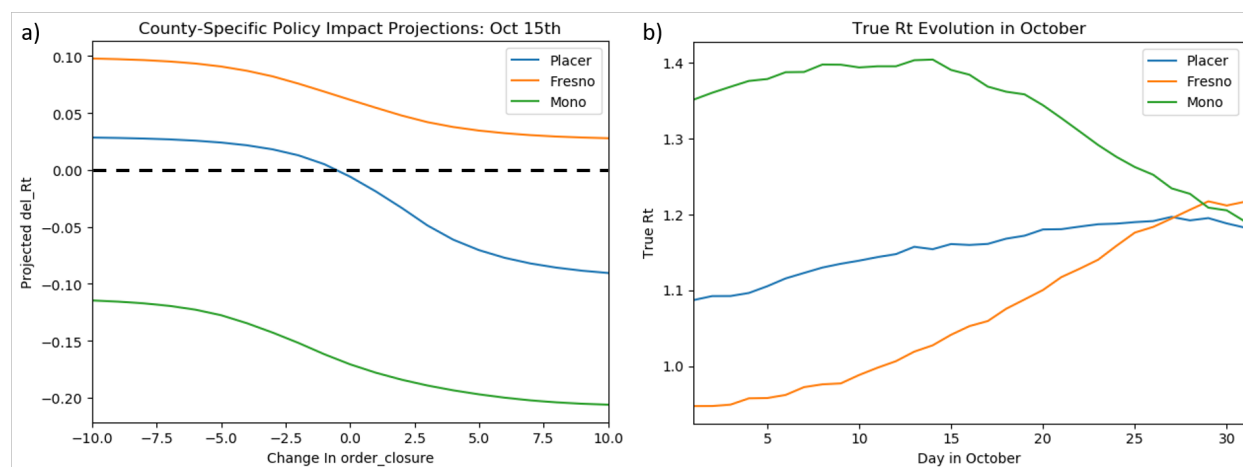


Figure 3: **Selected County Projections on October 15th.** a) Our model projects that Mono, Placer, and Fresno counties face drastically different policy outlooks on Oct 15th; Mono is on a nice downward trend, Placer is at an inflection point, and Fresno is in the midst of a surge that even maximal business closures couldn't fully reverse. In reality all three counties maintained loose closure policy for the entire period shown. b) These predictions are well reflected in the true  $R_t$  curves from that time.

## Competing Interests

All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: financial support from the Clinard Innovation Fund, which has no conflict of interest with this work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

## Acknowledgments

We thank Dr. Jeremy Goldhaber-Fiebert for providing us with early access to the SC-COSMO public health policy data set, which helped kick-start this research effort.

## References

- [1] Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. London: Imperial College COVID-19 Response Team, March. 2020; 16.
- [2] Fang Y, Nie Y, Penny M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *J Med Virol* 2020; 10.1002/jmv.25750
- [3] Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, Wu J. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *J Clin Med* 2020; 9: 462
- [4] Kennedy DM, Zambrano GJ, Wang Y, Neto OP. Modeling the effects of intervention strategies on COVID-19 transmission dynamics. *J Clin Vir* 2020; 128: e104440
- [5] Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, Priesemann V. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions *Science* 2020; 369(6500):eabb9789
- [6] Prem K, Liu Y, Russell T, Kucharski AJ, Eggo RM, Davies N. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Health* 2020;5(5):e261-e270
- [7] Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, et al. Inferring the effectiveness of government interventions against COVID-19. *Science* 2020; DOI:10.1126/science.abd9338
- [8] Fouda A, Mahmoudi N, Moy N, Paolucci F. The COVID-19 pandemic in Greece, Iceland, New Zealand, and Singapore: Health policies and lessons learned. *Health Policy and Technology* 2020; 9(4):510-524

- [9] Van Dyke ME, Rogers TM, Pevzner E, Satterwhite CL, Shah HB, Beckman WJ, Ahmed F, Hunt DC, Rule J. Trends in County-Level COVID-19 Incidence in Counties With and Without a Mask Mandate — Kansas, June 1–August 23, 2020. *Morb Mortal Wkly Rep* 2020; 69(47): 1777-1781
- [10] Dave D, Friedson AI, Matsuzawa K, Sabia JJ. When do Shelter-In-Place Orders Fight COVID-19 Best? Policy Heterogeneity Across States and Adoption Time. *Economic Inquiry* 2021; 59(1):29-52
- [11] New York Times Coronavirus tracking team. New York Coronavirus Map and Case Count. <https://www.nytimes.com/interactive/2020/us/new-york-coronavirus-cases.html>
- [12] Qian Z, Alaa A, van der Schaar M. When and How to Lift the Lockdown? Global COVID-19 Scenario Analysis and Policy Assessment using Compartmental Gaussian Processes. *NeurIPS* 2020; <https://arxiv.org/pdf/2005.08837.pdf>
- [13] T. Hale, A. Petherick, T. Phillips, and S. Webster. Variation in government responses to COVID-19. Blavatnik School of Government Working Paper, 2020.
- [14] Goldhaber-Fiebert JD, Holsinger A, Holsinger E, Long E. State- and County-Level COVID-19 Public Health Orders in California: Constructing a Dataset and Describing Their Timing, Content, and Stricture. *medRxiv* 2020; DOI: 10.1101/2020.11.08.20224915
- [15] Kevin Systrom, Thomas Vladek and Mike Krieger. Rt.live GitHub repository (2020): <https://github.com/rtcovidlive/covid-model>
- [16] <https://covidactnow.org/us/california-ca>
- [17] <https://www.google.com/covid19/mobility/>
- [18] Feng S, Feng Z, Ling C, Chang C, Feng Z. Prediction of the COVID-19 epidemic trends based on SEIR and AI models. *PLOS ONE* 2021; DOI: 10.1371/journal.pone.0245101
- [19] Distanto C, Pereira IG, Goncalves LMG, Piscitelli P, Miani A. Forecasting Covid-19 Outbreak Progression in Italian Regions: A model based on neural network training from Chinese data. *medRxiv* 2020; DOI: 10.1101/2020.04.09.20059055
- [20] Yang X, Zeng Z, Wang K, Wong S, Liang W, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J. Thorac Dis* 2020; 12(3): 165-174
- [21] Jo HC, Kim J, Huang TC, Ni YL. CONDLSTM-Q: A Novel Deep Learning Model for Predicting COVID-19 Mortality in Fine Geographical Scale. *arXiv* 2020; <https://arxiv.org/pdf/2011.11507.pdf>
- [22] Bathwal R, Chitta P, Tirumala K, Varadarajan V. Ensemble Machine Learning Methods for Modeling COVID19 Deaths. *arXiv* 2020; <https://arxiv.org/pdf/2010.04052.pdf>
- [23] Sill J, Abu-Mostafa YS. Monotonicity Hints. *NeurIPS* 1996; <https://authors.library.caltech.edu/64696/1/1270-monotonicity-hints.pdf>
- [24] Magdon-Ismail M, Chen H-C, Abu-Mostafa YS. The Multilevel Classification Problem and a Monotonicity Hint. *IDEAL* 2002: pp 410-415
- [25] Diuvesteijn W, Feelders A. Nearest Neighbour Classification with Monotonicity Constraints. *Machine Learning and Knowledge Discovery in Databases* 2008: pp 301-316
- [26] Kotlowski W, Slowinski R. Rule learning with monotonicity constraints. *ICML* 2009: pp 537-544, DOI: 10.1145/1553374.1553444
- [27] Wehenkel A, Louppe G. Unconstrained Monotonic Neural Networks. *NeurIPS* 2019; <https://arxiv.org/pdf/1908.05164.pdf>
- [28] Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. *arXiv* 2017; <https://arxiv.org/pdf/1412.6980.pdf>
- [29] <https://covid19forecasthub.org/>